

REMARKS

In the March 15, 2004 Office Action, the Examiner noted that claims 1-21 were pending in the application; rejected claims 1-4, 6, 10, 11, 14-16 and 18-21 under 35 U.S.C. § 102(e); and rejected claims 5, 7-9, 12, 13 and 17 under 35 U.S.C. § 103. In rejecting the claims, U.S. Patents 6,675,170 to Flake (Reference A in the March 15, 2004 Office Action); 6,078,913 to Aoki et al. (Reference A in the September 17, 2003 Office Action) were cited. Claims 6 and 9 have been cancelled and thus, claims 1-5, 7, 8 and 10-21 remain in the case. The Examiner's rejections are traversed below.

Newly Cited Prior Art: U.S. Patent 6,675,170 to Flake

The Flake patent is directed to partitioning large hyperlinked databases by a hyperlink structure into a subset 102 (Fig. 1) of relevant documents. Subset 102 is identified using "an s-t minimum cut/maximum flow network algorithm to perform a partitioning of a database ... [based on] pre-specified source and sink documents (which correspond, respectively, to desirable and undesirable documents)" (column 4, lines 49-53). The subset 102, or

cut set is a collection of links that have the property that removing them from the database completely disconnects the source documents from the sink documents. Hence, after identifying the cut set and removing those links from the database, all documents that can be reached by starting from a source document and following any remaining links belong to the set of desirable documents. All documents that cannot be reached from a source document through any link are undesirable.

(column 4, lines 55-63). The source documents are referred to as "one or more prototype documents 100" (column 4, lines 24-25, see Fig. 1) or "as 'seed' documents" (column 4, line 30), each of which "is chosen to be a very general example of a desired subset to be identified (e.g., a baseball WWW page with many links, or an index to universities), while the sink document ... is preferably chosen to be generic and representative of the entire database" (column 4, line 64 to column 5, line 2). The documents that are compared with the "seed" documents may be "the World Wide Web" (column 4, lines 4-5) represented by a database created by a "web crawler", such as the database maintained by "Yahoo!, Lycos, or Excite" (column 5, line 2). For this method to work, "the number of links coming in and out of the source and sink nodes must be greater in number than the size of the cut set. Thus, it may be necessary to use more than one prototype and combine them as a composite source or sink node until this criteria is met" (column 5, lines 12-17). The user is required to supply the seed document(s) (see column 5, lines 36-37).

Prior Art Rejections

In item 2 on pages 2-6 of the Office Action, claims 1-4, 6, 10, 11, 14-16 and 18-21 were rejected under 35 U.S.C. § 102(e) as anticipated by Flake and in item 4 on pages 6-8, claims 7-9, 12 and 13 were rejected under 35 U.S.C. § 103(a) as unpatentable over Flake. As discussed above, Flake identifies a subset of documents from a database to be used for subsequent searches, relying on one or more seed documents supplied by a user and generic sink document(s), then a minimum cut/maximum flow network algorithm is applied to identify the smallest “cut set” that includes all documents linked to the source document(s). The number of documents in the cut set is not predetermined in any manner. Thus, the method taught by Flake does not meet the limitation “collecting documents equal to or larger, in number, than a predetermined value from inside a community” (e.g., claim 1, lines 3-4) and then “collecting documents from inside and outside the community based on the reference of collected documents” (e.g., claim 1, lines 5-6).

As discussed above, claim 1 recites a significantly different operation than what is disclosed by Flake. To ensure that the smallest cut set has been properly identified, Flake has to determine that there are no links between any of the documents in the cut set and the remaining documents. In comparison, the present invention starts with an existing community of documents and collects documents within that community, up to a predetermined number “based on a reference” and then collects additional documents both “inside and outside the community based on the reference” (claim 1, line 5). This is a process that can be stopped at anytime, either by the system or the user, with information that may be useful to the user. The method taught by Flake is merely the first step in the process of finding relevant documents, by reducing the set of documents that will be searched. Since claims 19 and 21 recite limitations similar to claim 1 and claims 2-4 and 10-15 depend from claim 1, it is submitted that claims 1-4, 10-15, 19 and 21 patentably distinguish over Flake for the reasons discussed above.

In rejecting claim 11, it was asserted that column 4, line 16 of Flake “discloses assigning a keyword to the collected document[s] based on a referencing expression used in the collected document[s]” (Office Action, page 3, lines 19-20). However, the paragraph at column 4, lines 11-22 of Flake describes an “implicit link ... [as] a weighted value that expresses a relationship between two documents such as keyword overlap” (column 4, lines 14-16) and gives as an example “two documents [that] have identical keywords” (column 4, lines 16-17). It is apparent from the statements in the paragraph at column 4, lines 11-22 of Flake, that the keywords have already been assigned when they are compared and thus, there is nothing in this paragraph

teaching or suggesting “assigning a keyword to the collected documents based on a referencing expression used in the collected documents” (claim 11, lines 2-3). Therefore, it is submitted that claim 11 further patentably distinguish over Flake.

In the paragraphs at the end of page 6 and top of page 7, the Examiner acknowledged that Flake fails to disclose the limitations previously recited in claim 7, and instead discloses using a minimum cut/maximum flow network algorithm in determining whether documents are related. The Examiner asserted that the modification to Flake required to meet the limitations recited in claim 7 would have been obvious to one of ordinary skill in the art at the time the invention was made, but no evidence was cited to support this assertion. The Examiner is respectfully requested to either cite evidence in the prior art supporting the assertion in accordance with MPEP § 2144.03, or withdraw the rejection. Until such evidence is provided, it is submitted that claim 7 patentably distinguishes over Flake.

Similarly, on page 7 of the Office Action, the Examiner acknowledged that Flake fails to disclose the limitations recited in claim 8, but that it would be obvious to modify “Sato ‘517” (Office Action, page 7, line 16) to include the limitation (it is assumed that the quotation is a typographical error and that it was intended to say that it would be obvious to modify Flake). However, the only motivation for making the modification to Flake was “for the purpose of searching for documents based o[n] the most important two keywords” (Office Action, page 7, last line); once again without citing any evidence in the prior art. Therefore, for the same reasons discussed above with respect to claim 7, it is submitted that claim 8 patentably distinguishes over Flake unless evidence is cited supporting the rejection.

In rejecting claim 13, it was asserted that column 3, lines 30-35 of “Flake discloses counting a number of different documents referenced using the referencing expression” (Office Action, page 8, lines 12-13). Presumably, what is being referenced is the sentence that reads “[t]he source document preferably comprises a plurality of seed documents, each of which is of the desired type” (column 3, lines 30-31). While this sentence indicates that there are multiple documents, there is no suggestion of counting the number of them and more importantly there is no suggestion that these are documents “referenced using the referencing expression” (claim 13, line 2). As noted above, in the method taught by Flake the “seed documents” are supplied by the user. No suggestion was cited or has been found that the method taught by Flake is aware of any relationship between the seed documents and a referencing expression, or that any documents are counted which are referenced using a referencing expression. As a result, there can be no suggestion of “not assigning the keyword based on the referencing expression

when the number of different documents is equal to or larger than a predetermined value” (claim 13, last two lines). Furthermore, as in the case of claims 7 and 8, nothing was cited to support the assertion that a person of ordinary of skill in the art “would have been motivated to modify Flake ... for the purpose of improving the invention by determining a significant group of documents which are not covered by a keyword” (Office Action, page 8, last three lines) to meet the limitations recited in claim 13. For the above reasons, it is submitted that claim 13 further patentably distinguishes over Flake.

In rejecting claim 16, it was asserted that column 5, lines 1-5 of Flake discloses the limitations recited on the last three lines of claim 16. However, this portion of Flake describes the “sink” document as “generic and representative of the entire database” (column 5, lines 1-2) or “linked to all of the documents in the database or a random subset thereof” (column 5, lines 4-5). There is nothing in this portion of Flake that teaches or suggests “receiving the documents retrieved **separately** from inside and outside the community” (claim 16, line 4, emphasis added). As discussed above, there is no “community” of documents in the system taught by Flake from which documents are obtained separately from other documents. The method taught by Flake starts with source documents, sink documents and documents to be searched. There is no retrieval of documents from the source or sink documents, these are already in the possession of the system and there is no other partitioning of the remaining documents, until the “cut set” is formed. Therefore, it is submitted that claim 16 patentably distinguishes over Flake.

In rejecting claim 18, it was asserted that column 4, lines 23-32 discloses the limitations recited therein. However, as discussed above, this portion of Flake only describes the formation of the “cut set” as separate from the remaining documents. Claim 18 recites “determining a prospect to be collected next based on a reference between a positive sample document group which is a document group related to a field and a negative sample document group which is a document group less related to the field” (claim 18, lines 3-5). The “cut set” is created on the exact **opposite** of what is recited in claim 18: the documents in the cut set are selected based on a **lack** of a reference between the documents in the cut set and outside the cut set. Therefore, it is submitted that claim 18 patentably distinguishes over Flake.

In rejecting claim 20, it was asserted that column 3, lines 12-35 and column 4, lines 23-32 disclose the limitations recited therein. However, as discussed above, the cited portion of column 4 discusses creation of the cut set and the cited portion of column 3 is the portion of the Summary describing use of the maximum flow algorithm in identifying the cut set and applying a further search method to further partition the first subset or filtering out documents. However,

nothing has been found in either portion cited in the rejection of claim 20 that teaches or suggests “determining a document to be collected relating to the field based on a reference to the positive sample document group and the negative sample document group” (claim 20, lines 6-7, emphasis added). Therefore, it is submitted that claim 20 patentably distinguishes over Flake.

In item 5 on pages 9-10 of the Office Action, claims 5 and 17 were rejected under 35 USC § 103(a) as unpatentable over Flake in view of Aoki et al. First, it is submitted that Aoki et al. does not teach or suggest modifying Flake to overcome the deficiencies noted above with respect to claim 1 from which claim 5 depends. Second, it is unclear what in the cited portion (Fig. 1 and column 5, lines 12-35) of Aoki et al. is relevant to “determining whether or not ... [a] document is in the community according to information indicating the position of the document in the network” (claim 5, lines 2-3). All that is disclosed in Aoki et al. is a database storing contents and locations of documents. There is no suggestion that this information is used to determine anything and certainly not whether a document is in a “community” as that term is used in claim 1. Therefore, it is submitted that claim 5 further patentably distinguishes over Flake in view of Aoki et al. due to the additional limitation recited therein. If the Examiner is relying on an assertion of obviousness to modify Flake to use the information stored in the system disclosed by Aoki et al. the Examiner is requested to cite evidence supporting this assertion in accordance with MPEP § 2144.03, as discussed above with respect to claim 7.

The last two lines of claim 17 includes a limitation similar to that discussed above with respect to claim 1 as distinguishing over Flake taken alone. As noted above, Aoki et al. does not teach or suggest any modification of Flake to meet the limitation recited at the end of claim 17. Therefore, it is submitted that claim 17 patentably distinguishes over Flake in view of Aoki et al. for the reasons discussed above with respect to claim 1. In addition, lines 5-7 of claim 17 recites a limitation similar to that discussed above with respect to claim 5. Therefore, it is submitted that claim 7 further patentably distinguishes over Flake in view of Aoki et al. for the reasons discussed above with respect to claim 5.

Entry of Amendment

Claims 6 and 9 have been cancelled, claims 7 and 8 have been amended to form independent claims by incorporating the limitations of claim 6 and claims 10 and 11 have been amended to use language that conforms to the language used in claim 1 from which they depend. Therefore, no “new search” should be required to reconsider the amended claims. Entry of the Amendment and reconsideration of the claims is earnestly solicited.

Summary

It is submitted that the references cited by the Examiner, taken individually or in combination, do not teach or suggest the features of the present claimed invention. Thus, it is submitted that claims 1-5, 7, 8 and 10-21 are in a condition suitable for allowance.

Reconsideration of the claims and an early Notice of Allowance are earnestly solicited.

Finally, if there are any formal matters remaining after this response, the Examiner is requested to telephone the undersigned to attend to these matters.

If there are any additional fees associated with filing of this Amendment, please charge the same to our Deposit Account No. 19-3935.

Respectfully submitted,

STAAS & HALSEY LLP

Date: 6/15/04

By: Richard A. Gollhofer
Richard A. Gollhofer
Registration No. 31,106

1201 New York Avenue, NW, Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501